AD NO. ____________________________
DTC PROJECT NO. 8-CO-160-UXO-016
REPORT NO. ATC-9297

SHALLOW WATER

UXO TECHNOLOGY DEMONSTRATION SITE

SCORING RECORD NO. 3

SITE LOCATION:
U.S. ARMY ABERDEEN PROVING GROUND

DEMONSTRATOR:
IT JEWELL, INC.
11654 PLAZA AMERICA DRIVE NO. 631
RESTON, VA 20190-4700
TECHNOLOGY TYPE/PLATFORM:
MAGNETOMETER

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

JANUARY 2007

Prepared for:
U.S. ARMY ENVIRONMENTAL COMMAND
ABERDEEN PROVING GROUND, MD 21010-5401

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

DISTRIBUTION UNLIMITED, JANUARY 2007.

CSTE-DTC-AT-SL-E

MEMORANDUM FOR RECORD

SUBJECT: Operations Security (OPSEC) Review of Paper/Presentation

1. The attached record entitled "The Shallow Water UXO Technology Demonstration Site Scoring Record #3" dated December 2006 is provided for review for public disclosure in accordance with AR 530-1 as supplemented. The scoring record is proposed for public release via the internet.

2. I, the undersigned, am aware of the intelligence interest in open source publications and in the subject matter of the information I have reviewed for intelligence purposes. I certify that I have sufficient technical expertise in the subject matter of this report and that, to the best of my knowledge, the net benefit of this public release outweighs the potential damage to the essential secrecy of all related ATC, DTC, ATEC, Army or other DOD programs of which I am aware.

Gary Rowe	Gary Rowe	December 2006
NAME (Printed)	SIGNATURE	DATE

CONCURRENCE:	NAME (Printed)	SIGNATURE	DATE
Program Mgr/Customer (If not ATC owned technology)	Patrick McDonnell	Patrick McDonnell	18 DEC 06
Core Director	John Kopczynski	John Kopczynski	13Dec06.
Core OPSEC Tech POC	William Burch & Michael Karwatka	William Burch	30NOV06
ATC OPSEC Officer/ Security Manager	William Murr	Jenell Bigham	16 Jan 2007
Public Affairs Spec	Susan Hagan	Susan Hagan	24Jan07
Tech Dir, ATC	John R. Wallace	John R. Wallace	24Jan07

(Return to ATC PAO for further processing)

DTC GO/SES	N/A		

Encl
as

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
January 2007	Final	6 through 9 March 2006

4. TITLE AND SUBTITLE
STANDARDIZED UXO TECHNOLOGY DEMONSTRATION SITE SHALLOW WATER SCORING RECORD NO. 3 (IT JEWELL, INC.)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Rowe, Gary

5d. PROJECT NUMBER
8-CO-160-UXO-016

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Aberdeen Test Center
ATTN: CSTE-DTC-AT-SL-E
Aberdeen Proving Ground, MD 21005-5059

8. PERFORMING ORGANIZATION REPORT NUMBER
ATC-9297

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Environmental Command
ATTN: SFIM-AEC-ATT
Aberdeen Proving Ground, MD 21005-5401

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Same as Item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited.

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT
This report documents the efforts of IT Jewell, Inc. to detect and discriminate inert unexploded ordnance (UXO) using G882 and mm magnetometers. Testing was conducted at ATC, Standardized Shallow Water UXO Technology Demonstration Site. A description of the tested system and an estimate of survey costs along with the analysis of the system performance are provided.

15. SUBJECT TERMS
IT Jewell, Inc., UXO Standardized Technology Site, Shallow Water, magnetometer, MEC

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(Include area code)*
Unclassified	Unclassified	Unclassified	SAR		

ACKNOWLEDGMENTS

Author:

Gary W. Rowe
Military Environmental Technology Demonstration Center (METDC)
U.S. Army Aberdeen Test Center (ATC)
U.S. Army Aberdeen Proving Ground (APG)

Contributors:

William Burch
Military Environmental Technology Demonstration Center
U.S. Army Aberdeen Test Center
U.S. Army Aberdeen Proving Ground

Christina McClung
Aberdeen Data Services Team (ADST)
Logistics Engineering and Information Technology Company (Log. Sec/Tri-S)
U.S. Army Aberdeen Proving Ground

TABLE OF CONTENTS

SECTION 4. TECHNICAL PERFORMANCE RESULTS

PAGE

SECTION 5. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of munitions and explosives of concern (MEC), i.e., unexploded ordnance (UXO) and discarded military munitions (DMM), require testing so their performance can be characterized. To that end, the U.S. Army Aberdeen Test Center (ATC) located at Aberdeen Proving Ground (APG), Maryland, has developed a Standardized Shallow Water Test Site. This site provides a controlled environment containing varying water depths, multiple types of ordnance and clutter items, as well as navigational and detection challenges. Testing at this site is independently administered and analyzed by the government for the purposes of characterizing technologies, tracking performance during system development, and comparing the performance and costs of different systems.

The Standardized UXO Technology Demonstration Site Program is a multi-agency program spearheaded by the U.S. Army Environmental Command (USAEC). ATC and the U.S. Army Corps of Engineers Engineering, Research and Development Center (ERDC) provide programmatic support. The Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP), and the Army Environmental Quality Technology Program (EQT) provided funding and support for this program.

1.2 OBJECTIVE

The objective of the Shallow Water Standardized UXO Technology Demonstration Site is to evaluate the detection and discrimination capabilities of existing and emerging technologies and systems in a shallow water environment. Specifically:

a. To determine the demonstrator's ability to survey a shallow water area, analyze the survey data, and provide a prioritized "Target List" with associated confidence levels in a timely manner.

b. To determine both the detection and discrimination effectiveness under realistic scenarios that varies ordnance, clutter, and bathymetric conditions.

c. To determine cost, time, and manpower requirements needed to operate the technology.

1.3 CRITERIA

The scoring criteria specified in the Environmental Quality Technology - Operational Requirements Document (EQT-ORD) (app D, ref 1) for: A(1.6.a): UXO Screening, Detection and Discrimination document are presented in Table 1-1. Very little information was available on the capabilities of shallow water detection systems when these criteria were developed. However, they were used in the design of the test site, and the five metrics were used to measure system performance in this report.

TABLE 1-1. SCORING CRITERIA

Metric	Threshold	Objective
Detection	80% ordnance items buried to 1 foot and under 8 feet (2.4 m) of water at a standardized site detected	95% ordnance items buried to 4 feet and under 8 feet (2.4 m) of water at a standardized site detected
Discrimination	Rejection rate of 50% of emplaced non-UXO clutter at a standardized site with a maximum false negative rate of 10%	Rejection rate of 90% of emplaced non-UXO clutter at a standardized site with a maximum false negative rate of 0.5%
Reacquisition	Reacquire within 1 meter	Reacquire within 0.5 meter
Cost rate	$4000 per acre	$2000 per acre
Production rate	5 acres per day	50 acres per day

The ATC shallow water site is designed to evaluate the threshold-detection level of a range of ordnance at the 1-foot + 8-foot requirement. Limited information is available at the objective-detection level. All other measured results in this test were evaluated against both criteria levels.

1.4 APG SHALLOW WATER SITE INFORMATION

1.4.1 Location

The Aberdeen Area of APG is located in the northeast portion of Maryland on the western shore of the Chesapeake Bay in Harford County. The Shallow Water Test Site is located within a controlled range area of APG.

1.4.2 Soil Type

The area chosen for the shallow water test site was known as Cell No. 3 in a dredge-spoil field. The cell bottom is composed primarily of sediment removed from the Bush River. This is a freshwater site.

1.4.3 Test Areas

a. The test site contains five areas: calibration grid, blind test grid, littoral, open water, and deeper water. Additional detail on each area is presented in Table 1-2. A schematic of the calibration lanes is shown in Figure 1.

TABLE 1-2. TEST AREAS

Area	Description
Calibration grid	The calibration area contains 15 projectiles, 3 each 40, 60, 81, 105, and 155 mm. One of each projectile type is buried at the projectile diameter to depth ratio shown in Figure 1. This area is designed to provide the user with a sensor library of detection responses for the emplaced targets and an understanding of their resistivity prior to entering the blind test fields. Two "clutter-cloud" target scenarios have been constructed adjacent to this area (fig. 1).
Blind grid	The blind grid contains 644 detection opportunities. Each grid cell is 2 by 2 m^2. At the center of each cell is either an ordnance item, clutter, or nothing. Surrounding the blind grid on three sides are 3.6-kg (8-lb) shot puts, buried 0.3-meters deep in the sediment. The shot puts can be used as a navigational/ Global Positioning System (GPS) check. The GPS coordinates for the center of each grid and the shot put locations are provided to the vendor prior to testing.
Littoral	This is a sloping area on one side of the pond with vegetation growing into the water line. Water depth ranges from 0.3 to 1.8 meters. It contains a variety of navigational and detection challenges.
Open water	The open water scenario contains a variety of navigational, detection, and discrimination challenges. Water depth varies from 1.8 to 3.4 meters.
Deeper water	The water depth in this area varies between 3.4 and 4.3 meters.

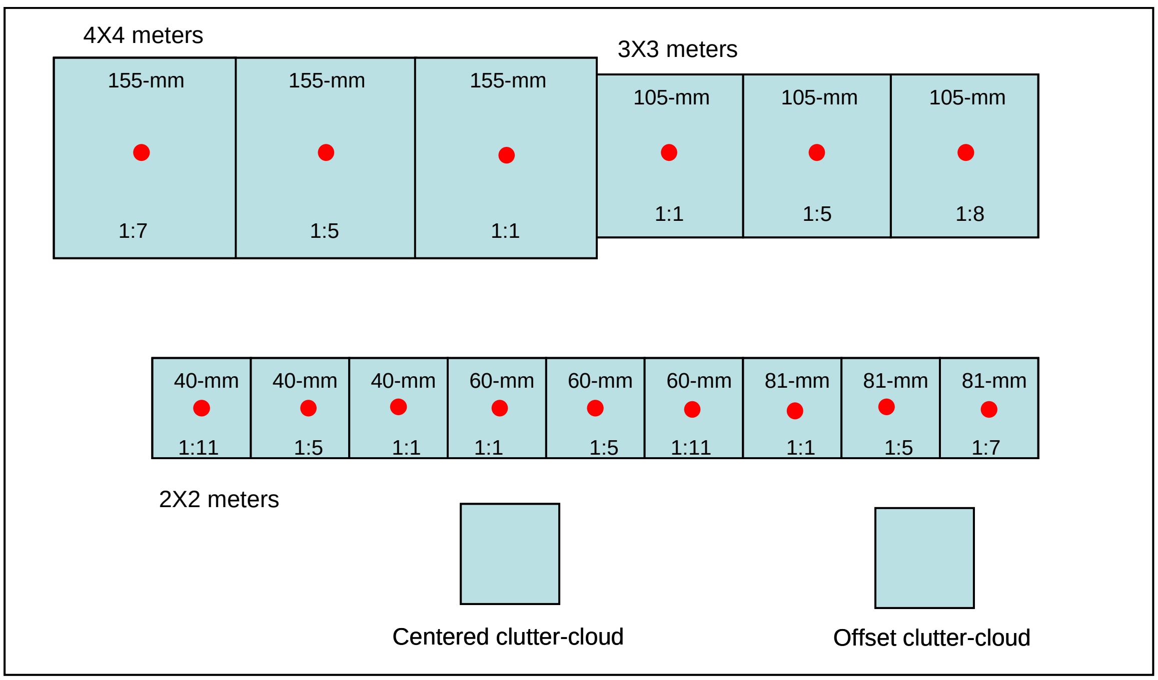

Figure 1. Schematic of the calibration grid.

b. The water depth at this facility during testing is maintained such that the calibration and blind grid areas meet the 2.4 meters (8 ft) detection criterion specified in section 1.3. The test site is approximately 2.8 hectares (6.9 acres) in size.

1.5 GROUND TRUTH TARGETS

The ground truth is composed of both inert ordnance and clutter items. The inert ordnance items are listed in Table 1-3. All items were located in storage sites at APG. The items have not been fired or degaussed.

Clutter items fit into one of three categories: ferrous, nonferrous, and mixed metals. The ferrous and nonferrous items have been further divided into three weight zones as presented in Table 1-4, and distributed throughout all test areas. Most of this clutter is composed of ordnance components; however, industrial scrap metal and cultural items are present as well. The mixed-metals clutter is composed of scrap ordnance items or fragments that have both a ferrous and nonferrous component and could reasonably be encountered in a range area. The mixed-metals clutter was placed in the open water area only.

TABLE 1-3. INERT ORDNANCE TARGETS

Description	**Length, mm**	**Diameter, mm**	**Aspect Ratio, W/L**	**Weight, *g***
40-mm L70 projectile	208	40	0.1923	965
60-mm mortar M49A2	185	60	0.3243	975
81-mm mortar M374	528	81	0.1534	3969
81-mm mortar M821	510	81	0.1588	3338
105-mm projectile M1	445	105	0.2360	13834
155-mm M107 projectile	684	155	0.2266	41731
8-in. M104/106	856	203	0.2371	89811

TABLE 1-4. CLUTTER WEIGHT RANGES

Clutter Type	**Weight Range in Grams**		
	Small	**Medium**	**Large**
Ferrous	10 to 510	511 to 2200	> 2201
Nonferrous	10 to 270	275 to 800	> 801

SECTION 2. SYSTEM UNDER TEST

2.1 DEMONSTRATOR INFORMATION

a. IT Jewell provided the information in sections 2.1 through 2.6 as part of either their Broad Agency Announcement proposal (app D, ref 2), or taken from the final report that was included with their data submittal (app D, ref 3). Section 2.8 contains ATC's comments on the demonstrated system.

b. This proposal was unique in that it provided two proven technologies to be run in back-to-back demonstrations, allowing the Army to evaluate and compare each technology on an even playing field by taking out the differences in operator experience, sensor platforms, and site conditions. This allowed for an objective comparison of the capabilities and limitations of each technology in the shallow water marine environment

c. IT Jewell is an independent systems integrator, surveyor, and data processor with no vested interest in one system's performing better than the other. Quality assurance and Quality Control (QA/QC) measures ensured that factors such as site conditions and operator experience were consistent between the two demonstrations to provide for an objective evaluation of system performance. Where applicable, the same navigation, platform, and data processing equipment and components were used to maintain the even playing field.

2.2 SYSTEM DESCRIPTION

The complete IT Jewell survey system consists of one of two sensors (described below), GPS navigation system, custom floating platform, low magnetic signature tow vessel, and water column depth sensor. The GPS system used is a Trimble Ag GPS 132 with integrated light bar. The platform is custom made from PVC plastic piping designed to securely mount the sensor at a predetermined depth below the water. The GPS antenna is mounted on the platform directly over the sensor mounting position with a vertical separation distance of approximately 1 meter to minimize potential sensor to antenna interference. The GPS A+B Swath pattern is set at 3 meters. The tow vessel is a plastic Pond Prowler pontoon boat powered by a mini-motor. The water column depth sensor is the ODOM Hydrotrac Echo. Magnetometer sensor and position data are captured through the MagLog™ software supplied by Geometrics (manufacturer of the G882 magnetometer).

Figure 2.   G882 cesium-based marine magnetometer.

Size and Weight: Body 2.75 in (7 cm) diameter, 4.5 ft (1.37 m) long with fin assembly (11 in cross width), 40 lbs (18 kg) includes sensor and electronics and one main weight.

Operating Principle: Self-oscillating split-beam Cesium Vapor (nonradioactive).

Swath: 2 m left and 2 m right yielding total swathing of 4.0 m.

Operating Range: 20,000 to 100,000 nT.

Counter Sensitivity: <0.004 nT/pHz rms.

Highest Sensitivity: 0.004 nT.

Output: RS-232 at 1,200 to 19,200 Baud.

Survey Frequency: 10 Hz.

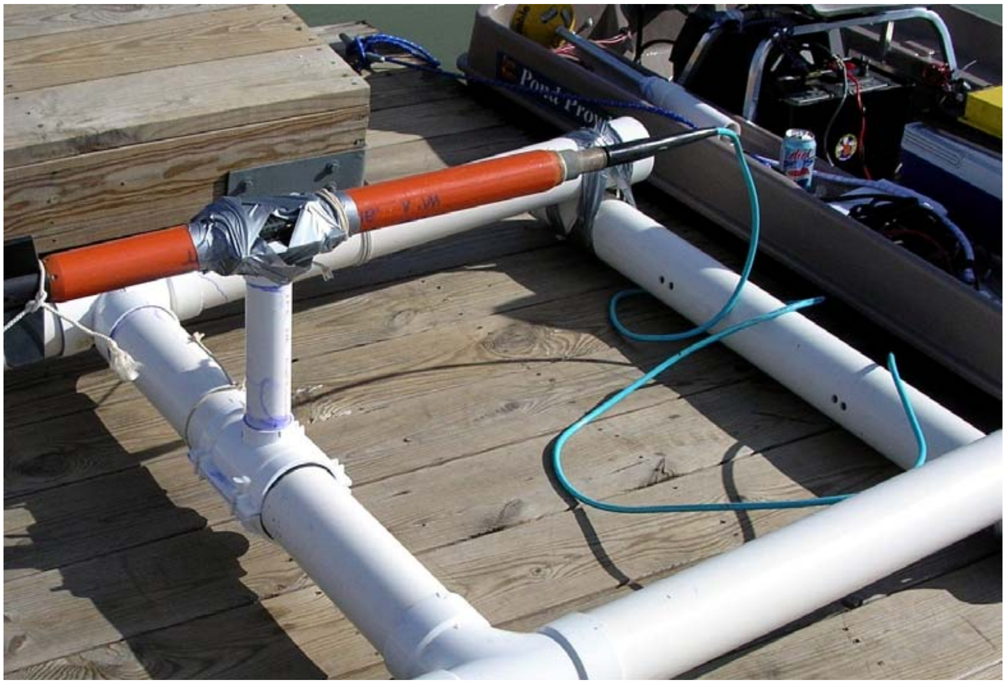

Figure 3. MM Explorer mini proton-based magnetometer.

Size and Weight: Body 2.375 in. (6 cm) diameter, 33.75 in. (86 cm) long with fin assembly, 8 lb (3.7 kg).

High sensitivity omni-directional Overhauser sensor.

Swath: 1.25 m left and 1.25 m right yielding total swathing of 2.5 m.

Operating Range: 18,000 to 120,000 nT.

Counter Sensitivity: 0.001 nT.

Sensor Sensitivity: 0.02 nT.

Output: RS-232, 9600 bps.

Survey Frequency: 1 Hz.

Figure 4. IT Jewell's shallow water UXO detection platform.

2.3 DEMONSTRATOR'S POC AND ADDRESS

POC: Mr. Shawn P. Jewell
email: sj@itjewell.com

Address: IT Jewell, Inc.
11654 Plaza America Drive No. 631
Reston, VA 20190-4700
(202) 744-0490

2.4 DEMONSTRATOR'S SITE SURVEY METHOD

a. Several runs or scenarios were done on both the G882 and MM Explorer Mini sensors. The complete pond run covered the entire pond. The perimeter run was along the perimeter of the pond. The spiral pond was a circular spiral run. The perimeter runs and spiral curve runs were done because preliminary data processing during the week of the survey showed some magnetic noise when the boat was turning. The perimeter and spiral curve runs surveys did not involve abrupt 180° turns, thereby providing data with less magnetic noise. Post-processing of the data combined the data sets to provide the optimal sensor responses.

b. The ODOM Hydrotrac Echo sounder was connected directly to the laptop through a RS232 port. The sensor head was mounted in the same cradle that held the magnetometers approximately 15 inches below the waterline. The data were processed to complete a bathymetric picture of the pond depth. These files were used to extract the depth values for the 644 points in the blind grid area as well as the targets found in the rest of the pond.

2.5 DEMONSTRATOR'S QC AND QA

a. Equipment Check Phase - This phase was a QC step during which the systems were rigorously tested and checked well in advance of the scheduled demonstration so that the inevitable system problems could be resolved prior to the actual survey. Access to the APG survey site was not needed for the predemonstration equipment check. The equipment check was done outside Army property in an area with similar characteristics.

b. Post-Survey QA Check - Before each survey system was disconnected for packaging and shipping, a scan of the data was done as a QA check. Any irregularities in the data were explained or resolved. If needed, portions of the test site were surveyed again to resolve any outstanding data issues.

2.6 DATA PROCESSING DESCRIPTION

a. Response Stage - Data on the boat turnarounds were analyzed and typically removed. Noise was evident as "steps," "shifts," and increased amplitude noise. Total Magnetic Intensity was then reviewed for noise by utilizing the fourth-difference (a 4th horizontal derivative developed as an indication of noise content in magnetic data). Suitable convolution filters were used to "smooth" the data. A 5-point Hanning window was used on the G-882 (0.5 seconds), and a 3-point Hanning window (3 seconds) was used on the MM Explorer Mini data. The data were then manually corrected for diurnal variation by fitting a first-order polynomial to the data to remove any first-order tilting. Finally, the data were plotted and reviewed for each sensor.

b. Discrimination Stage - Magnetic data were filtered using a High Pass 1D FFT filter algorithm to remove geologic signals. Signals with wavelengths longer than 10 meters were rejected. The data wavelengths were then picked using the calibration points provided by the Army. The filter was designed to enhance the signal seen over the calibration range, for application to the rest of the data. The Analytic signal was calculated from the processed data. This was calculated as the sum of the gradients and moved the peak Analytic signal over the magnetic source. This was also used as the ultimate discrimination value. Anomalies were manually picked from individual profiles while reviewing the Total Intensity, High Pass data, and Analytic signal. These were used to help calibrate the final picks. Targets were then picked from the Analytic signal grid. Targets below noise thresholds derived from the above calculations were rejected. Clutter was also largely included in this. Target positions, Total Magnetic Intensity (TMI), and Analytic signal (discrimination value) were tabled. Targets were finally discriminated, based on the range of observed data seen over the calibration range for each sensor individually.

2.7 DEMONSTRATOR'S SITE PERSONNEL

Project Geophysicist: Mr. Shawn Jewell

Data Acquisition Specialists: Mr. Peter Cooke

2.8 ATC'S SURVEY COMMENTS

a. This very basic, one-sensor system was easy and inexpensive to transport, set up, and operate. Positioning the magnetometers approximately 0.3 meters below the surface and using a shallow-draft boat allowed this system to maneuver and survey in most of the littoral zone. However, the depth of the sensor became a detriment to the system in the deeper water areas.

b. In interest of evaluating as many shallow water detection systems as possible, and realizing the opportunity to have the data from two different sensors collected by and processed using one source, this approach had the potential to yield meaningful site and sensor specific data. Many of the test variables needed for a valid side-by-side comparison appeared controlled to an accepted level.

SECTION 3. SURVEY COST ANALYSIS

3.1 DATES OF SURVEY

The IT Jewell magnetometer systems were tested from 6 through 9 March 2006.

Reconfiguring IT Jewell's system for the different magnetometers was as simple as inverting the floating platform, changing the sensor, reconnecting a cable, and returning the platform to its survey position. This allowed IT Jewell to survey different site areas with different sensors on the same day. For the purpose of this cost analysis, the areas surveyed by each sensor were summed and the operating costs were divided by the total area. This approach did not account for the extra time needed to calibrate two instead of one magnetometer and the downtime spent switching sensors. However, the percentage of time devoted to these activities was too small to warrant a separate cost analysis. Readers having specific interest in this information should refer to the daily log sheets in Appendix B.

3.2 SITE CONDITIONS

3.2.1 Atmospheric Conditions

An ATC weather station located adjacent to the test site recorded the average temperature and precipitation on an hourly basis for each day of operation. The temperatures listed in Table 3-1 represent the average temperature from 0700 through 1700. The hourly weather logs used to generate this summary are provided in Appendix A.

3.2.2 Water Conditions

Water conditions were monitored using a TIDALITE IV Portable Tide Gauge System©. Data recorded included water depth and temperature, significant wave height based on the average 1/3 wave height seen over the test period using the Draper/Tucker analysis method, and the full-wave frequency calculated by full-wave mean crossing detection. The values displayed in Table 3-1 were averaged from 0700 through 1700. Detailed information is provided in Appendix B.

TABLE 3-1. SITE CONDITION SUMMARY

Date, 06	Air Temperature, °C	Wind, km/h	Water Temperature, °C	Water Depth, m[a]	Significant Wave Height, m	Wave Frequency, Hz
6 Mar	5.1	7.7	6.8	0.3	0.0	0.2
7 Mar	4.5	15.7	Lost	Lost	Lost	Lost
8 Mar	4.4	7.3	6.8	0.3	0.0	0.2
9 Mar	12.4	11.0	6.7	0.4	0.0	0.3

[a]Variance between the required 2.4-meter test depth and actual test conditions.
Lost - instrumentation malfunction.

3.3 SURVEY ACTIVITIES

The information contained in this section provides an estimate of the time needed and costs associated with surveying an area with this demonstrator's system. This includes data on equipment setup and calibration, site survey and any resurvey time, and downtime due to system malfunctions and maintenance requirements.

3.3.1 Survey Times

a. A government representative monitored and recorded all on-site activities, which were grouped into one of 11 categories. The first eight categories were chargeable to the system while the last three were not. Categorizing these activities provided insight into the technical and logistical aspects of the system. The times recorded in each category were then matched with the number of demonstrator personnel, assigned skill levels, and a consistent (across-vendor) salary to produce an estimate of the survey costs.

(1) Initial setup/mobilization. Started at the time the demonstrator's equipment arrived at the survey site and stopped when the system was ready to acquire data.

(2) Daily setup/close-up. Monitored time spent mounting and dismounting the equipment each day.

(3) Instrument calibration. Recorded the amount of time used for daily quality assurance checks (e.g., sensors, GPS data, survey data quality).

(4) Data collection. Time spent surveying the test area.

(5) Downtime (nonsurvey time) for equipment/data checks. Covered time spent troubleshooting equipment or verifying survey tracks.

(6) Downtime (nonsurvey time) for equipment failure. Examples include replacing damaged cables, lost communication with base station, and any other failure that prevented surveying. Some weather-related failures fall into this category, for example, light-emitting diode (LED) displays darkened by the sun, wind creating waves too high to permit surveying, etc.

(7) Downtime (nonsurvey time) for maintenance. Battery replacement and memory downloads are typical examples.

(8) Demobilization. Commenced once the demonstrator completed the survey and concluded the final on-site check of the test data, and ended when the equipment and personnel were ready to leave the site.

(9) Nonchargeable downtime for breaks and lunch. The demonstrator's company policy set this standard.

(10) Nonchargeable downtime for weather-related causes (i.e., lighting, high wet-bulb heat index, and similar events).

(11) Nonchargeable downtime due to ATC range operating requirements. Danger zone conflicts, lack of support personnel, equipment or other ATC-caused delays.

b. Appendix C contains the daily log sheets. Table 3-2 summarizes that information to provide insight into the operational, maintenance, and logistic aspects of the system.

TABLE 3-2. TIME ON-SITE

Date (2006)	**6 Mar**	**7 Mar**	**8 Mar**	**9 Mar**	**Activity Totals, hr**
Activity (daily times recorded in minutes)					
Initial setup	120	-	330	-	18.7
Daily setup/close-up	155	125	30	120	12.6
Instrumentation calibration	-	30	100	60	3.2
Data collection	35	380	-	410	30.9
Equipment/data checks	-	60	-	10	1.2
Equipment failure	-	50	-	-	2.0
Maintenance	-	20	-	-	0.5
Demobilization	-	-	-	-	3.7
Breaks and lunch	-	-	-	-	0.0
Weather-related	-	-	-	-	0.0
ATC downtime	25	-	-	-	0.0
Daily total, hr	11.3	10.6	7.7	10.0	

Note: Task times have been rounded to 5-minute increments.

3.3.2 On-Site Data Collection Costs

The times associated with the 11 activities have been grouped into the three basic components of the evaluation: initial setup, site survey, and pack-up (demobilization). Note that site survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime for equipment/data checks or maintenance, downtime due to failure, and downtime due to weather. This combines the actual survey cost with the demonstrator’s associated on-site overhead costs.

A standardized estimate for labor costs associated with this effort was then calculated using the following job categories: supervisor ($95.00/hr), data analyst ($57.00/hr), and site support ($28.50/hr). The estimated costs are presented in Table 3-3.

TABLE 3-3. CALCULATED SURVEY COSTS

	No. of Persons	Hourly Wage	Hours	Cost
Initial Setup				
Supervisor	1	$95.00	18.7	$1,776.50
Data Analyst	1	$57.00	18.7	$1,065.90
Site Support	1	$28.50	18.7	$532.95
Subtotal				**$3,375.35**
Site Survey				
Supervisor	1	$95.00	50.4	$4,788.00
Data Analyst	1	$57.00	50.4	$2,872.80
Site Support	1	$28.50	50.4	$1,436.40
Subtotal				**$9,097.20**
Demobilization				
Supervisor	1	$95.00	3.7	$351.50
Data Analyst	1	$57.00	3.7	$210.90
Site Support	1	$28.50	3.7	$105.45
Subtotal				**$667.85**
Total on-site costs				**$13,104.40**

3.4 COST ANALYSIS

The data collection process described above provided an on-site cost guide to compare the performance of this vendor with any other that has demonstrated at the shallow water site. It is not a true indicator of survey costs. Many other expenses have not been included, such as travel costs, per diem, off-site data processing and analysis, company overhead, and profit.

Calculating the area surveyed was done by plotting the raw GPS coordinates and then combining the sensor swath (line spacing and associated overlap).

To determine the number of acres surveyed per day, the total number of hours spent at the test site (table 3-2) was divided by 8 (converts to 8-hr days). The number of acres was then divided by the number of 8-hour days. The cost per acre was determined by dividing the total survey costs (table 3-3) by the same number of acres. This information is summarized in Table 3-4.

TABLE 3-4. SURVEY COSTS

Area surveyed (acre[a])	7.34
Time on-site (8-hr days)	4.43
Calculated survey cost (U.S. dollars)	$5,884.30
Acres per day	1.66
Cost per acre	$801.68

[a]Acre = 4047 meters^2.

Table 3-5 presents a comparison of Tetra Tech's survey costs with the EQT-ORD criteria.

TABLE 3-5. TEST RESULTS - CRITERIA COMPARISION

Metric	**Threshold**	**Objective**	**IT Jewell**
Cost rate	$4000 per acre	$2000 per acre	$802 per acre
Production rate	5 acres per day	50 acres per day	1.66 acres per day

SECTION 4. TECHNICAL PERFORMANCE RESULTS

The technical evaluation of both magnetometers that IT Jewell used at the shallow water site is contained in this section. The analytical procedures are identical for each system.

4.1 AREA SURVEYED

4.1.1 Calculated Area

a. Both the test and scoring methodologies required the demonstrator to survey 100 percent of each of the four test areas (blind grid, open water, littoral, and deeper water). Scoring a partially surveyed area alters the ordnance and clutter sample sizes, and test area boundaries, and decreases the statistical confidence in the performance statements made for that area. Allowing partial scoring decreases the validity of performance comparisons made between multiple test areas for a single demonstrator and comparisons made between multiple demonstrators for a single test area.

b. Realizing that some systems may not be able to survey 100 percent of a given test area, a ranking system was established. The percent coverage for a given test area is determined by first plotting the raw GPS coordinates combined with the sensor swath (line spacing and associated overlap), calculating the area surveyed, and then comparing the surveyed area with the total test area.

$$\frac{\text{Section Surveyed}}{\text{Test Area Size}} \times 100 = \%\ \text{Surveyed}$$

c. The demonstrator's system is always scored against the complete ground truth for a given test area regardless of the percentage covered.

4.1.2 Area Assessment

The ranking system and survey results are presented in Table 4-1.

TABLE 4.1. M 882 SURVEY RANKING SYSTEM AND RESULTS

Ranking System		Survey Results			
% Area Covered	**Ranking**	**Test Area**	**G882 % Area Covered**	**MM % Area Covered**	**Data Use**
95 to 100	Met				Direct comparison between systems and areas.
90 to 94	Generally met				Comparison between systems and areas. A small negative bias is contained in the reported numbers (bias not quantified in this report).
50 to 89	Partially met	Blind grid	78	81	Reported, not compared between systems or areas. A large negative bias is contained in the reported numbers (bias not quantified in this report).
		Open water	69	58	
		Deeper water	72	58	
		Littoral	64	70	
0 to 49	Not met				Not scored/not reported.

4.2 SYSTEM SCORING PROCEDURES

a. The scoring entities used in this program were predicated on knowing the composition and location of every detectable item in an area. The deeper water area is the one exception. Ground-truth targets were placed in this area without a pre-survey and clearing operation. Therefore, only the system's probability of detection (P_d) was evaluated in this area.

b. The best indicator of survey performance is the blind grid. This area provides a statically valid, controlled environment in which the demonstrator must provide a response (ordnance, clutter, or blank) at each of the 644 locations. Comparison of the response and discrimination lists to the ground truth in this area both determines the range of ordnance the system can reliably detect and establishes the baseline to which system performance in all other test areas is measured.

c. The scoring terms and definitions, along with an explanation of the receiver operating characteristics (ROC) curve development and the chi-square analysis used in this report, are provided in Appendix C.

d. Demonstrator performance was scored in two stages: response and discrimination.

e. Response stage scoring evaluates the ability of the demonstrator's system to detect emplaced ground-truth targets without regard to discriminating ordnance from clutter. In this stage, the GPS locations and signal strengths of all anomalies the demonstrator deemed sufficient for further investigation and/or processing are reported. This list was generated with minimal processing, i.e., associating signal strength with GPS location, and includes only signals that are above the system noise level.

f. The discrimination stage evaluated the demonstrator's ability to segregate ordnance from clutter. The same GPS locations reported in the response stage anomaly list were evaluated on the basis of the demonstrator's discrimination process (section 2.6). A discrimination stage list was generated and prioritized on the basis of the demonstrator's determination that an anomaly was more likely to be ordnance rather than clutter. Typically, higher output values indicate a higher confidence that an ordnance item is present at a specified location. The demonstrator then specifies the threshold value for the prioritized ranking that provides optimal system performance. This value is the discrimination stage threshold.

g. Both the response and discrimination lists contain the identical number of potential target locations, differing only in the priority ranking of the declarations.

h. Within both of these stages, the following entities were measured:

(1) P_d.

(2) Probability of false positive (P_{fp}).

(3) Probability of background alarm (P_{ba})/background alarm rate (BAR).

4.2.1 <u>ROC Curves</u>

a. Based on the entire range of ground truth targets used at this site, ROC curves were generated for both the response and discrimination stages. In both stages, the probability of detection verses false alarm rates was plotted. False alarms were divided into two groups: (1) anomalies corresponding to emplaced clutter items, thereby measuring the P_{fp}, and (2) anomalies not corresponding to any known item, termed background alarms (P_{ba}) in the blind grid area and background alarm rate (BAR) in all other areas.

b. The ROC curves for the response and discrimination stages for all areas surveyed are shown in Figures 5 through 16. Horizontal lines illustrate the system performance at the demonstrator's recommended noise level during the response stage, or discrimination threshold level in the discrimination stage. The point where the curve crosses the horizontal line defines the subset of targets the demonstrator recommends digging.

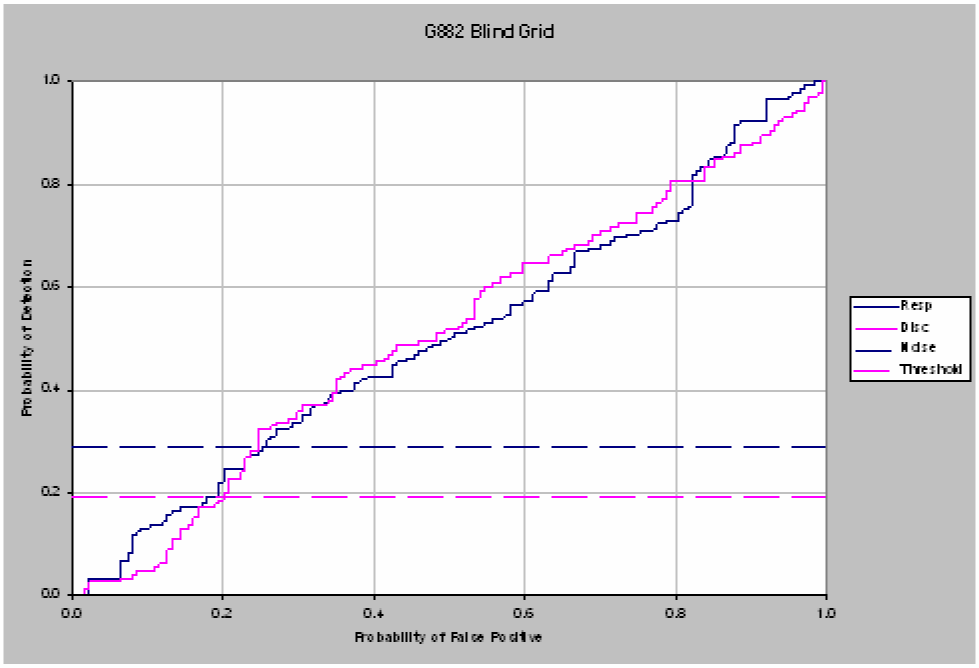

Figure 5. G882 magnetometer - blind grid P_d versus P_{fp}.

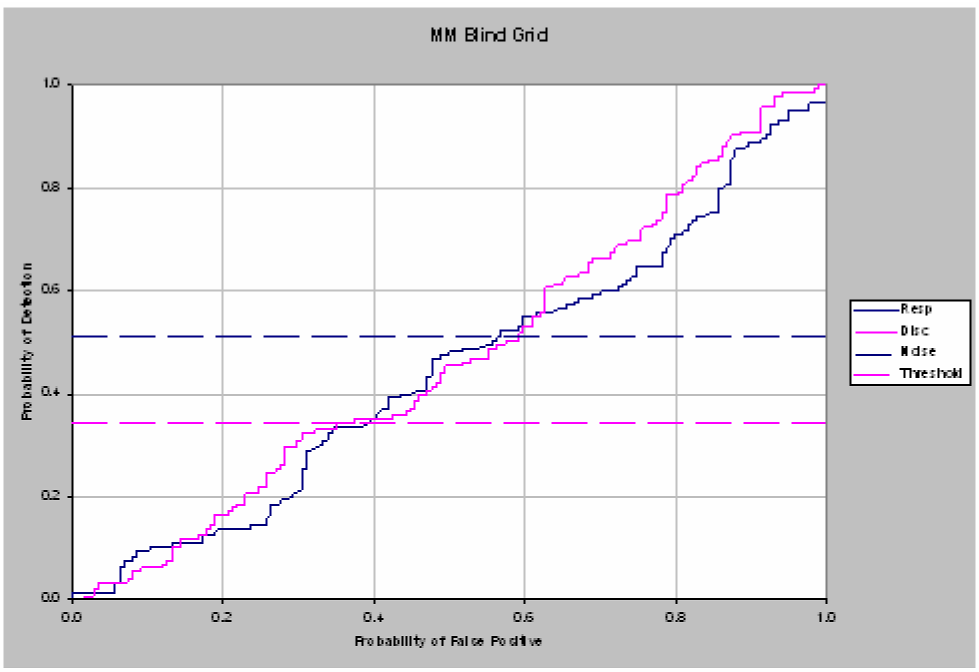

Figure 6. MM magnetometer - blind grid P_d versus P_{fp}.

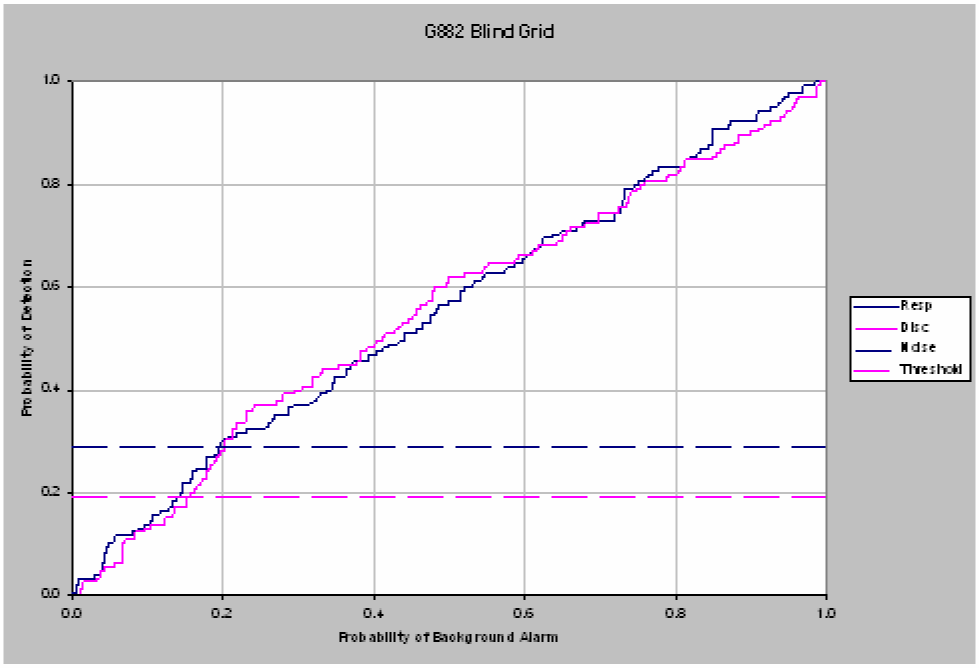

Figure 7. G882 magnetometer - blind grid P_d versus P_{ba}.

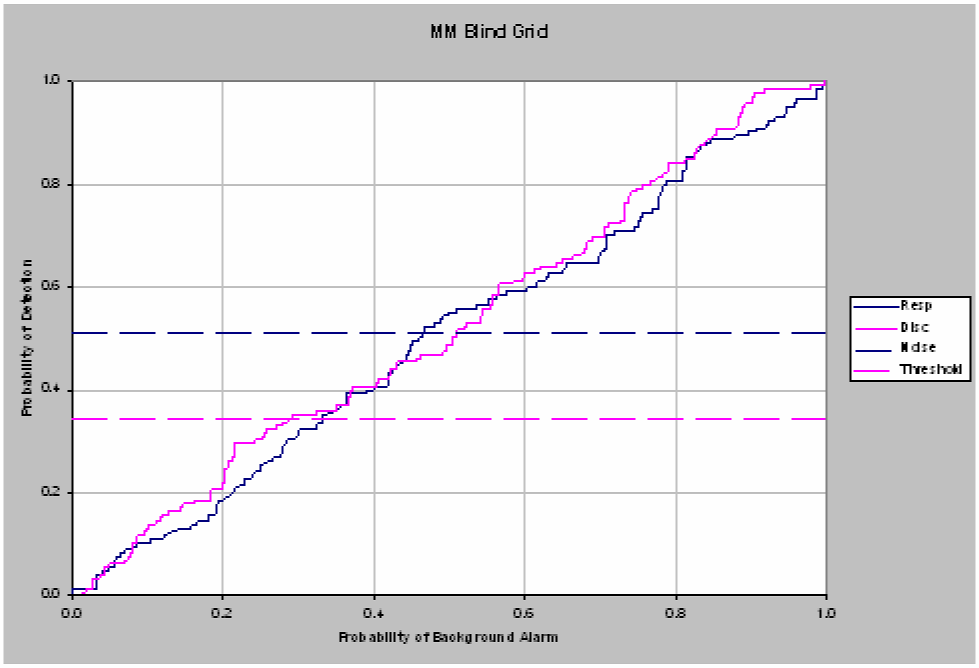

Figure 8. MM magnetometer - blind grid P_d versus P_{ba}.

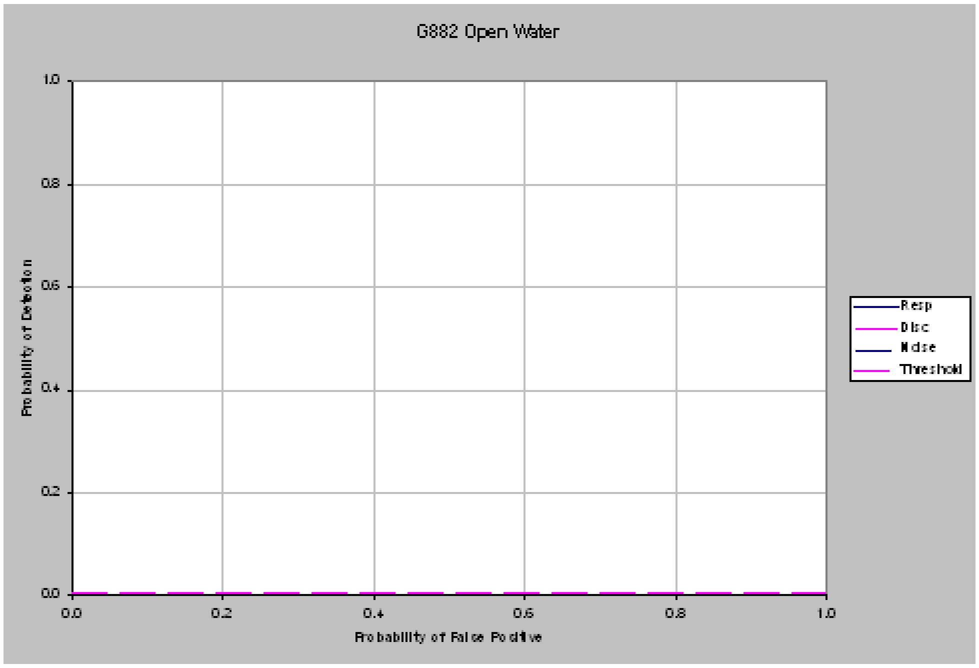

Figure 9. G882 magnetometer - open water P_d versus P_{fp}.

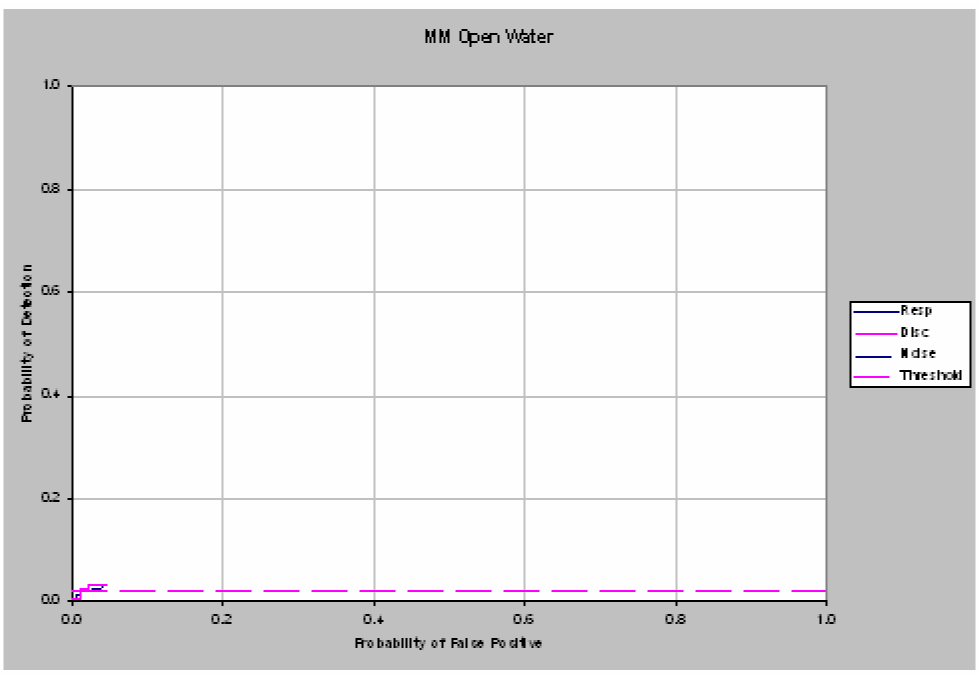

Figure 10. MM magnetometer - open water P_d versus P_{fp}.

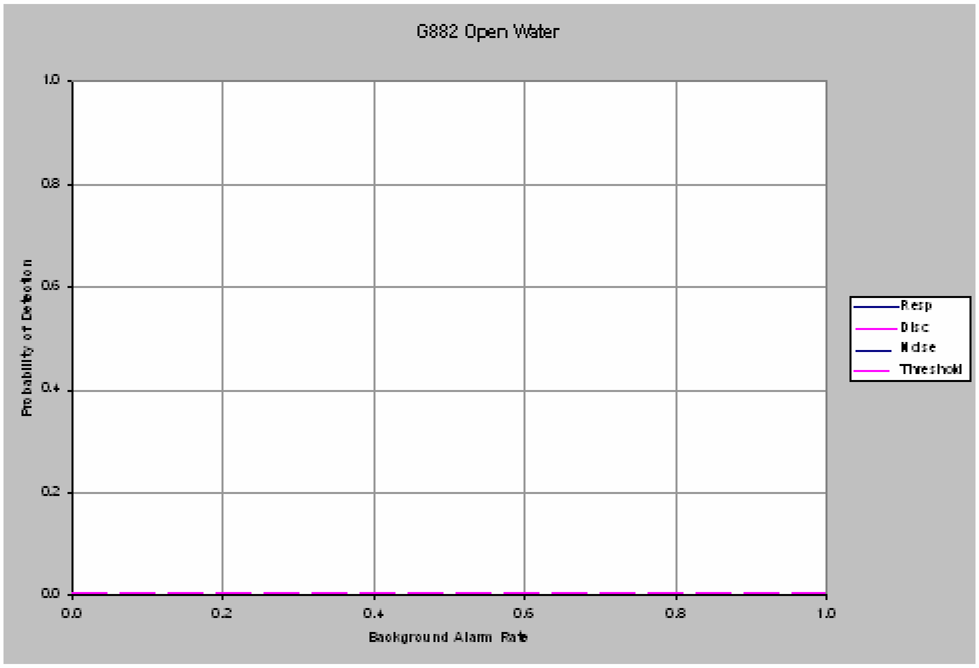

Figure 11. G882 magnetometer - open water P_d versus BAR.

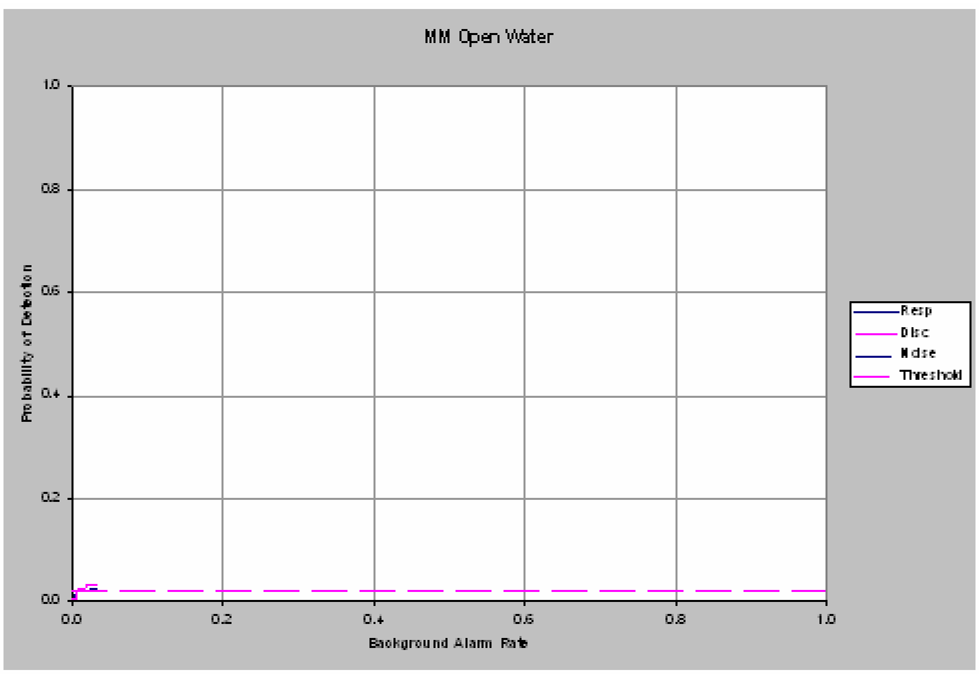

Figure 12. MM magnetometer - open water P_d versus BAR.

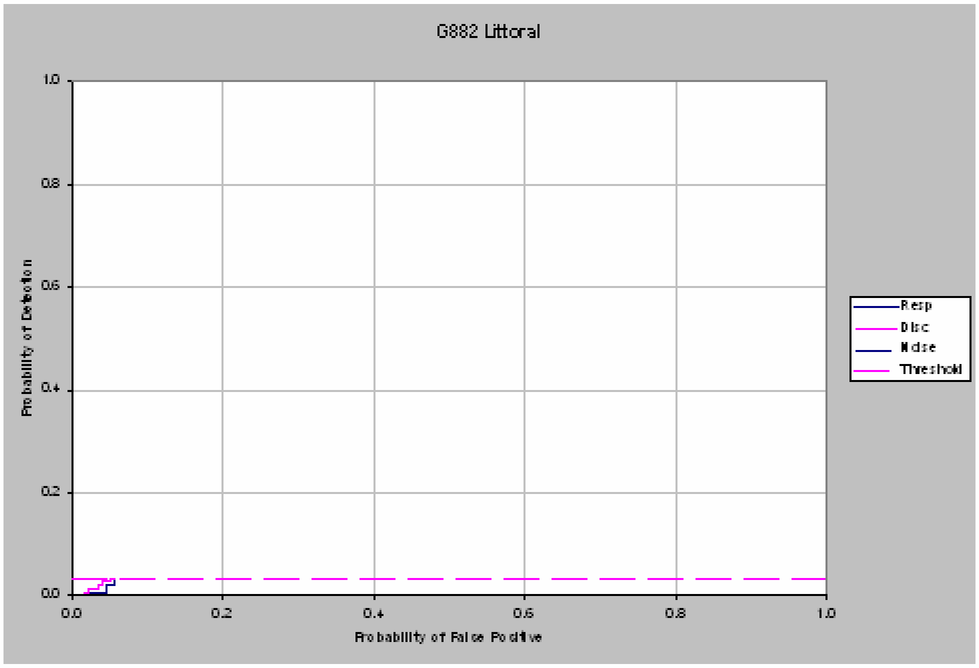

Figure 13. G882 magnetometer - littoral P_d versus P_{fp}.

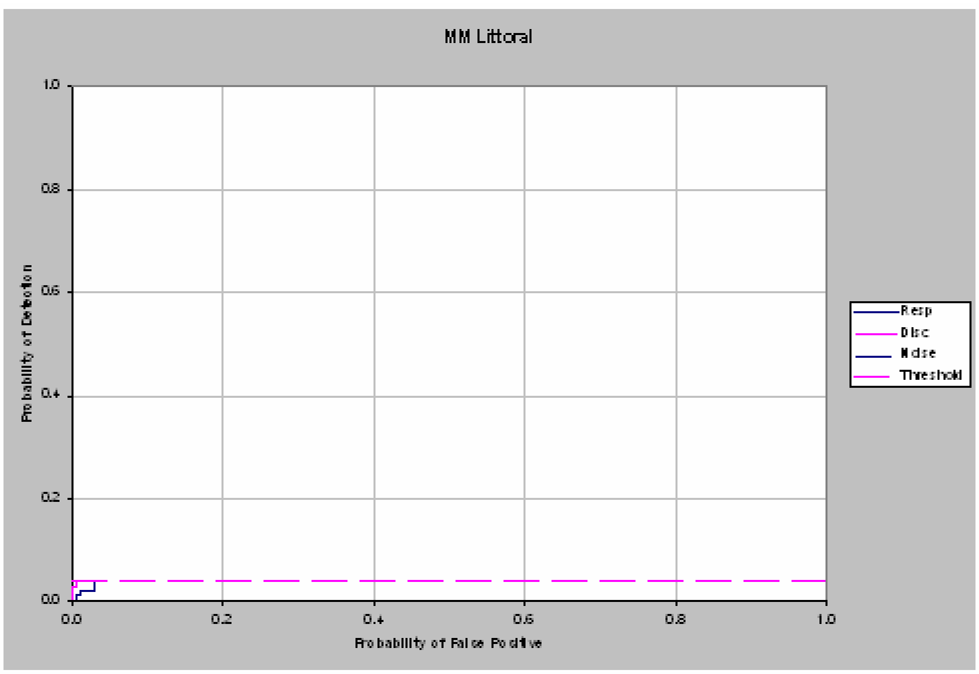

Figure 14. MM magnetometer - littoral P_d versus P_{fp}.

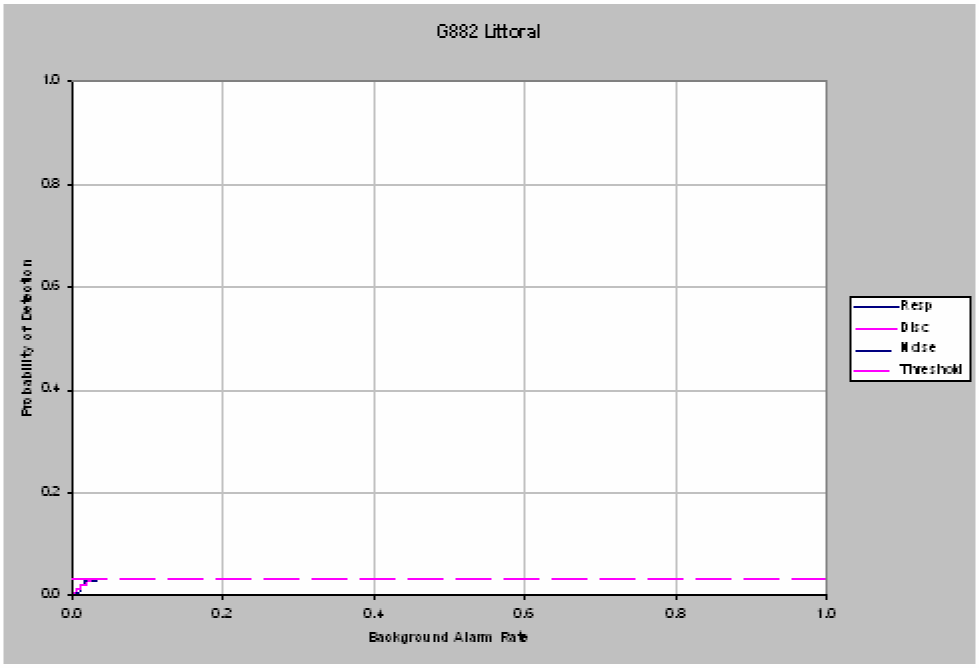

Figure 15. G882 magnetometer - littoral P_d versus BAR.

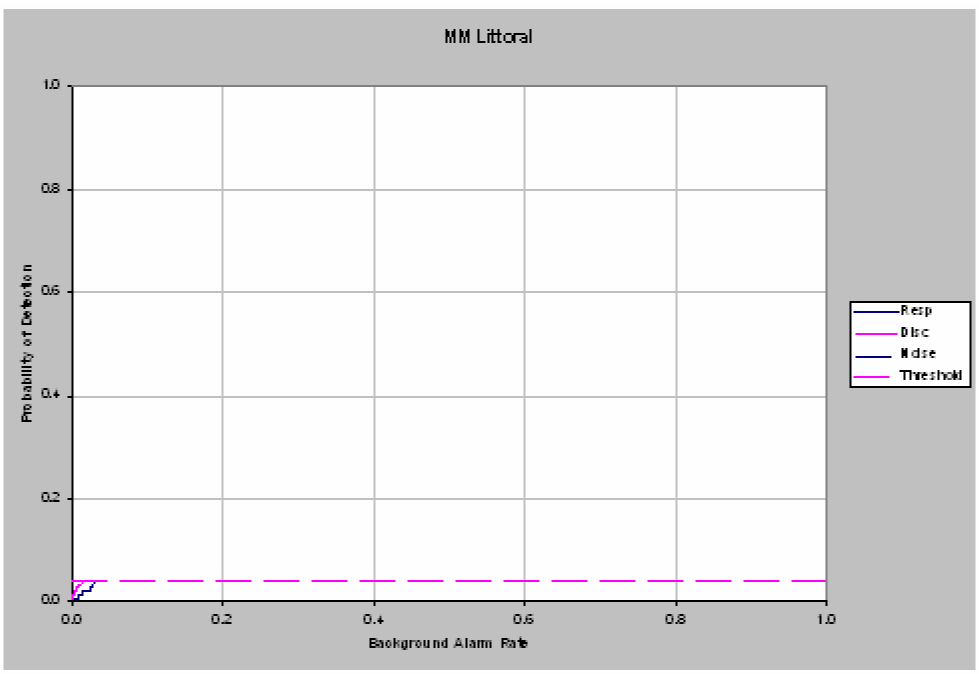

Figure 16. G882 magnetometer - littoral P_d versus BAR.

4.2.2 Detection Results

Detection results, broken out by stage, area surveyed, and ordnance size, are presented in Table 4-2. The results by size indicate how well the demonstrator detected/discriminated ordnance of a given caliber. Overall results summarize ordnance detection over a given area. All values were calculated assuming the number of detections was a binomially distributed random variable. These results are reported at the 90-percent reliability/95-percent confidence levels unless otherwise noted.

TABLE 4-2. G882 SYSTEM DETECTION SUMMARY

Metric	Overall	By Projectile Caliber					
		40 mm	60 mm	81 mm	105 mm	155 mm	8 in.
Blind grid							
Response stage							
P_d	29.0%	24.1%	24.1%	31.0%	31.0%	34.5%	
P_d lower 90% confidence	24.0%	14.0%	14.0%	19.7%	19.7%	22.6%	
P_{fp}	25.3%						
P_{fp} lower 90% confidence	21.0%						
P_{ba}	19.7%						
Discrimination stage							
P_d	19.3%	20.7%	13.8%	6.9%	24.1%	31.0%	
P_d lower 90% confidence	15.1%	11.2%	6.2%	1.8%	14.0%	19.7%	
P_{fp}	20.1%						
P_{fp} lower 90% confidence	16.2%						
P_{ba}	15.1%						
Open water							
Response stage							
P_d	0.6%	0.0%	0.0%	3.4%	0.0%	0.0%	0.0%
P_d lower 90% confidence	0.1%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%
P_{fp}	2.5%						
P_{fp} lower 90% confidence	1.2%						
BAR m^{-2}	0.022						
Discrimination stage							
P_d	0.6%	0.0%	0.0%	3.4%	0.0%	0.0%	0.0%
P_d lower 90% confidence	0.1%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%
P_{fp}	2.5%						
P_{fp} lower 90% confidence	1.2%						
BAR m^{-2}	0.021						
Littoral region							
Response stage							
P_d	3.4%	13.8%	0.0%	3.4%	0.0%	0.0%	
P_d lower 90% confidence	1.7%	6.2%	0.0%	0.4%	0.0%	0.0%	
P_{fp}	5.7%						
P_{fp} lower 90% confidence	3.6%						
BAR m^{-2}	0.038						
Discrimination stage							
P_d	3.4%	13.8%	0.0%	3.4%	0.0%	0.0%	
P_d lower 90% confidence	1.7%	6.2%	0.0%	0.4%	0.0%	0.0%	
P_{fp}	5.7%						
P_{fp} lower 90% confidence	3.6%						
BAR m^{-2}	0.038						
Deeper water							
Response stage							
P_d	0.0%					0.0%	
P_d lower 90% confidence	0.0%					0.0%	
Discrimination stage							
P_d	0.0%					0.0%	
P_d lower 90% confidence	0.0%					0.0%	
Response stage noise level: 52625							
Recommended discrimination threshold: 4.18							

TABLE 4-3. MM SYSTEM DETECTION SUMMARY

Metric	Overall	By Projectile Caliber					
		40 mm	60 mm	81 mm	105 mm	155 mm	8 in.
Blind grid							
Response stage							
P_d	51.0%	51.7%	48.3%	48.3%	44.8%	62.1%	
P_d lower 90% confidence	45.4%	38.4%	35.1%	35.1%	31.9%	48.5%	
P_{fp}	56.3%						
P_{fp} lower 90% confidence	51.2%						
P_{ba}	46.5%						
Discrimination stage							
P_d	34.5%	34.5%	24.1%	31.0%	27.6%	55.2%	
P_d lower 90% confidence	29.3%	22.6%	14.0%	19.7%	16.8%	41.7%	
P_{fp}	37.4%						
P_{fp} lower 90% confidence	32.5%						
P_{ba}	29.2%						
Open water							
Response stage							
P_d	1.9%	3.4%	3.4%	0.0%	3.4%	0.0%	0.0%
P_d lower 90% confidence	0.7%	0.4%	0.4%	0.0%	0.4%	0.0%	0.0%
P_{fp}	2.5%						
P_{fp} lower 90% confidence	1.2%						
BAR m^{-2}	0.022						
Discrimination stage							
P_d	1.9%	3.4%	3.4%	0.0%	3.4%	0.0%	0.0%
P_d lower 90% confidence	0.7%	0.4%	0.4%	0.0%	0.4%	0.0%	0.0%
P_{fp}	2.5%						
P_{fp} lower 90% confidence	1.2%						
BAR m^{-2}	0.021						
Littoral region							
Response stage							
P_d	4.1%	6.9%	0.0%	3.4%	0.0%	10.3%	
P_d lower 90% confidence	2.2%	1.8%	0.0%	0.4%	0.0%	3.9%	
P_{fp}	2.9%						
P_{fp} lower 90% confidence	1.4%						
BAR m^{-2}	0.029						
Discrimination stage							
P_d	4.1%	6.9%	0.0%	3.4%	0.0%	10.3%	
P_d lower 90% confidence	2.2%	1.8%	0.0%	0.4%	0.0%	3.9%	
P_{fp}	2.9%						
P_{fp} lower 90% confidence	1.4%						
BAR m^{-2}	0.029						
Deeper water							
Response stage							
P_d	0.0%					0.0%	
P_d lower 90% confidence	0.0%					0.0%	
Discrimination stage							
P_d	0.0%					0.0%	
P_d lower 90% confidence	0.0%					0.0%	
Response stage noise level: 52620							
Recommended discrimination threshold: 5.099							

4.2.3 System Discrimination

Using the demonstrator's recommended setting, the items detected and correctly classified as ordnance were further evaluated as to whether the demonstrator could correctly identify the ordnance type. The list of ground-truth ordnance items was provided to the demonstrator prior to testing.

IT Jewell's "dig-list" discriminated between ordnance and clutter, but not between ordnance types. The latter was an optional requirement.

4.2.4 System Effectiveness

Efficiency and rejection rates were calculated to quantify the discrimination ability at two specific points of interest on the ROC curve: the point where no decrease in P_d occurred (i.e., the efficiency is by definition equal to one) and the operator-selected threshold. These values, for both magnetometers, are presented in Table 4-3.

TABLE 4-3. EFFICIENCY AND REJECTION RATES

	Efficiency		**False-Positive Rejection Rate**		**Background Alarm Rejection Rate**	
Blind Grid						
	G882	**MM**	**G882**	**MM**	**G882**	**MM**
At operating point	0.67	0.68	0.20	0.34	0.23	0.37
With no loss of P_d	1.00	1.00	0.00	0.00	0.03	0.00
Open Water						
	G882	**MM**	**G882**	**MM**	**G882**	**MM**
At operating point	1.00	1.00	0.00	0.00	0.06	0.04
With no loss of P_d	1.00	1.00	0.40	0.20	0.65	0.56
Littoral Region						
	G882	**MM**	**G882**	**MM**	**G882**	**MM**
At operating point	1.00	1.00	0.00	0.00	0.00	0.00
With no loss of P_d	1.00	1.00	0.10	0.80	0.37	0.55

4.2.5 Chi-Square Analysis

Typically, a chi-square 2 by 2 Contingency Test for comparison between ratios is used to compare performance across test areas with regard to P_d^{res}, P_d^{disc}, P_{fp}^{res}, and P_{fp}^{disc}, efficiency, and false alarm rejection rates. The intent of the comparison is to determine if the features introduced in each test site have a degrading effect on the performance of the sensor system.

Neither the G882 nor MM magnetometer surveys covered enough of any test area to permit a valid comparison of performance, either between areas or between sensor types.

4.2.6 Location Accuracy

An insufficient quantity of data points (coordinates of ordnance items in any area that were first detected in the response stage within a 0.5-meter radius of their true positions, then correctly identified as ordnance in the discrimination stage) exists to draw any conclusions on system location accuracy for either of the two magnetometers tested.

The comparison between the results obtained during testing and the EQT-ORD criteria is presented in Tables 4-5 and 4.6.

TABLE 4-5. G882 TEST RESULTS - CRITERIA COMPARISION

Metric	**Threshold**	**Objective**	**G882 by Area**	
Detection	80% ordnance items buried to 1 foot and under 8 feet (2.4 m) of water.	95% ordnance items buried to 4 feet and under 8 feet (2.4 m) of water.	Blind grid	29 %
			Open water	0.6 %
			Littoral	3.4 %
Discrimination	Rejection rate of 50% of emplaced non-UXO clutter.	Rejection rate of 90% of emplaced non-UXO clutter.	Blind grid	20 %
			Open water	0 %
			Littoral	0 %
	Maximum false negative rate of 10%.	Maximum false-negative rate of 0.5%.	Not assessed. An analytical procedure is not available to address this criterion.	
Reacquisition	Reacquire within 1 meter.	Reacquire within 0.5 meter.	The number of correctly identified items is insufficient to draw any conclusions.	

Note: The blind grid and open water areas are in general accordance with the threshold requirements.

TABLE 4-6. MM TEST RESULTS - CRITERIA COMPARISION

Metric	**Threshold**	**Objective**	**MM by Area**	
Detection	80% ordnance items buried to 1 foot and under 8 feet (2.4 m) of water.	95% ordnance items buried to 4 feet and under 8 feet (2.4 m) of water.	Blind grid	51 %
			Open water	1.9 %
			Littoral	4.1 %
Discrimination	Rejection rate of 50% of emplaced non-UXO clutter.	Rejection rate of 90% of emplaced non-UXO clutter.	Blind grid	34 %
			Open water	0 %
			Littoral	0 %
	Maximum false negative rate of 10%.	Maximum false-negative rate of 0.5%.	Not assessed. An analytical procedure is not available to address this criterion.	

TABLE 4-6 (CONT'D)

Metric	Threshold	Objective	MM by Area
Reacquisition	Reacquire within 1 m	Reacquire within 0.5 m	The number of correctly identified items is insufficient to draw any conclusions.

Note: The blind grid and open water areas are in general accordance with the threshold requirements.

SECTION 5. APPENDIXES

APPENDIX A. TEST CONDITIONS LOG

ATMOSPHERIC CONDITIONS

Date, 06	Time, EDT	Average Wind Direction, deg	Average Wind Speed, km/h	Wind Direction Average Standard Deviation, deg	Peak Wind Speed, km/h	Average Temperature, °C
6 Mar	0700	17	3.7	27	6.3	-1.8
	0800	341	3.7	17	6.8	0.2
	0900	345	6.8	24	14.6	3.5
	1000	7	9.0	26	18.5	5.2
	1100	342	10.8	24	19.0	5.6
	1200	354	7.2	34	16.6	7.1
	1300	124	8.4	32	15.4	7.7
	1400	169	12.1	35	21.2	7.4
	1500	194	9.0	15	14.6	7.3
	1600	190	6.0	18	10.1	7.5
	1700	228	7.6	28	11.3	6.4
7 Mar	0700	354	14.8	14	24.6	-0.5
	0800	353	16.1	14	29.0	-0.4
	0900	353	17.9	14	30.1	0.8
	1000	348	15.9	17	25.6	2.5
	1100	338	14.0	19	24.5	3.8
	1200	347	15.3	19	31.7	5.5
	1300	353	15.4	21	27.4	6.4
	1400	3	15.0	18	27.5	7.2
	1500	344	16.6	20	32.5	7.8
	1600	340	17.4	14	28.6	8.2
	1700	356	14.8	19	28.5	8.2
8 Mar	0700	22	2.3	82	5.0	-5.1
	0800	350	1.6	98	5.8	-2.2
	0900	309	2.4	24	5.1	1.3
	1000	343	3.4	30	9.7	3.7
	1100	230	5.8	52	16.4	5.6
	1200	189	9.8	28	18.5	6.1
	1300	184	11.6	29	21.1	7.4
	1400	192	12.9	17	20.9	8.1
	1500	194	13.4	11	19.1	7.9
	1600	192	8.4	11	15.3	7.7
	1700	192	8.7	15	13.7	8.0
9 Mar	0700	126	3.4	16	9.8	5.8
	0800	128	1.8	48	5.3	7.3
	0900	133	1.6	29	8.0	9.2
	1000	94	2.3	87	5.3	9.8
	1100	169	11.9	33	27.8	11.7
	1200	190	18.3	12	29.8	13.7
	1300	186	18.2	13	29.3	15.1
	1400	187	19.6	11	31.2	15.7
	1500	189	16.7	12	28.2	16.0
	1600	181	15.6	10	27.0	16.0
	1700	174	11.4	13	19.5	16.4

The water conditions during the IT Jewell survey are shown in Figures A-1 through A-3.

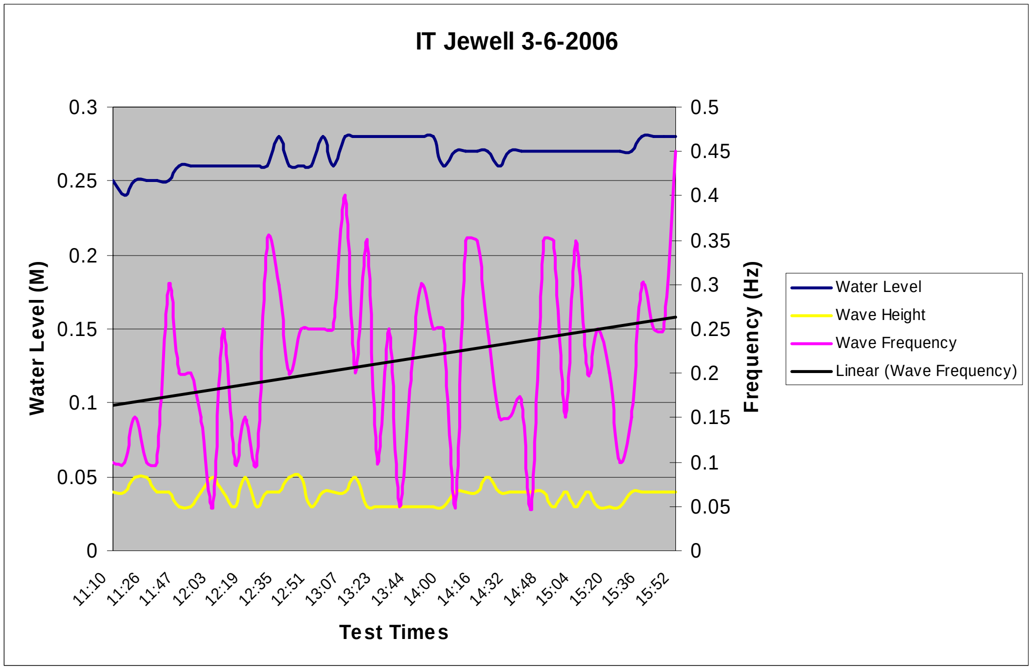

Figure A-1. Water conditions on 6 March 2006.

Water condition results for 7 March 2006 were lost because of an instrument malfunction.

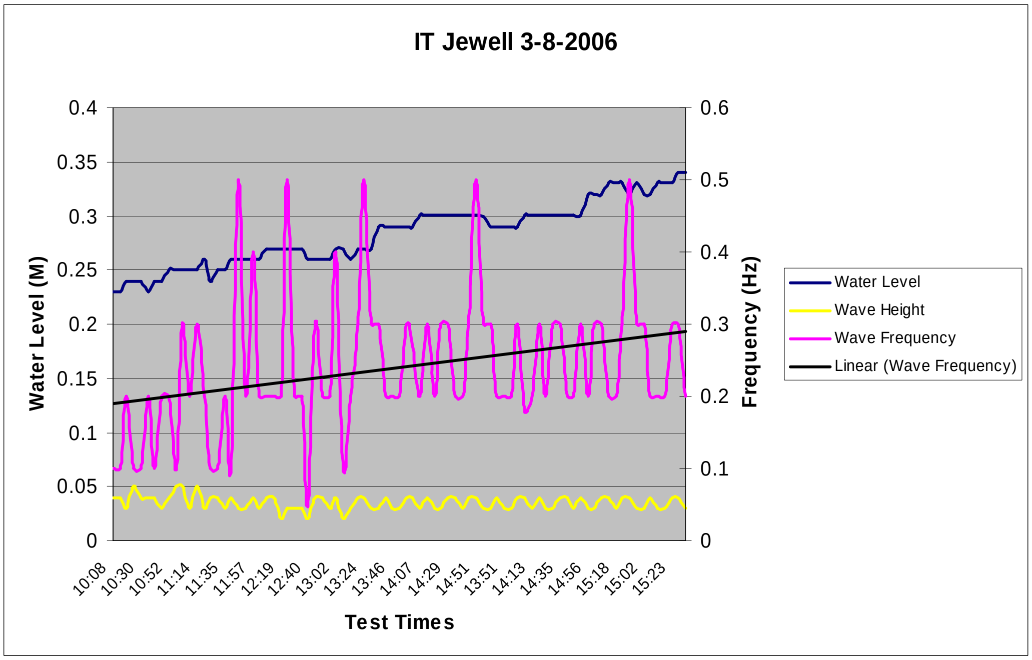

Figure A-2. Water conditions on 8 March 2006.

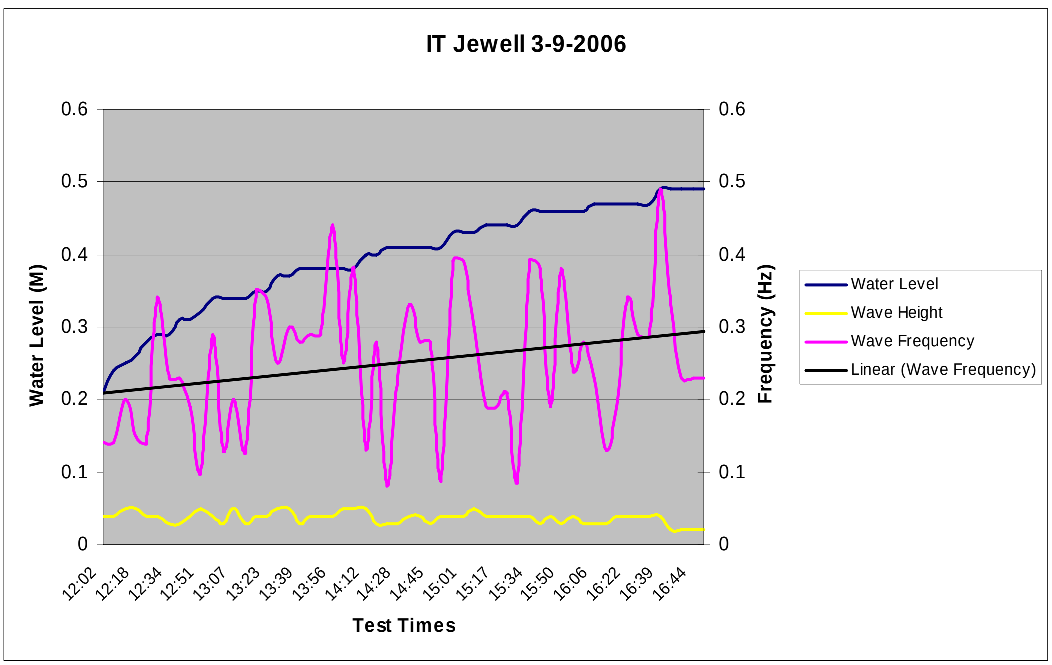

Figure A-3. Water conditions on 9 March 2006.

APPENDIX B. DAILY ACTIVITIES LOG

Company: IT Jewell Date: 6 March 2006			**Personnel: Shawn Jewell, Peter Cooke**	
Start	**Stop**	**Remarks**	**Activity**	**Chargeable, min**
0925	0945	Arrived at test site. Safety briefing/questions	Downtime (ATC)	25
0945	1105	Unloaded the truck. Launched the boat. Attached the mini proton-based magnetometer to the mounting platform.	Initial setup	120
1105	1140	Using the boat and boat paddle, the water depth around the perimeter of the pond was measured	Data collection	35
1140	1215	Lunch	Nonchargeable downtime	35
1215	1325	Attached sensor platform to the boat. Installed associated instrumentation, GPS, data collection computer, batteries.	Initial setup	70
		Note: The plan was to survey one area with one sensor, then survey the same area with the second sensor. Sensor depth was 18 in. for both probes, in all areas		
1325	1400	Surveyed the calibration lanes using the mini-probe	Data collection	35
1400	1440	Switched to Geometrics 882 sensor	Setup	40
1440	1520	Surveyed the calibration lanes using the G-882 magnetometer	Data collection	40
1520	1545	Pond perimeter survey with the G-882 sensor	Data collection	25
1545	1600	End of day cleanup	Daily close-up	15

Company: IT Jewell Date: 7 March 2006			**Personnel: Shawn Jewell, Peter Cooke**	
Start	**Stop**	**Remarks**	**Activity**	**Chargeable, min**
0730	0920	ATC delay - monthly safety meeting	Downtime (ATC)	110
0920	1130	Set up	Daily setup	140
1130	1205	Perimeter survey using G-882 sensor	Data collection	35
1205	1245	Lunch	Nonchargeable downtime	40
1245	1640	Surveying using a criss-cross pattern, minor axis of pond.	Data collection	235
1600	1620	Replaced motor battery	Downtime equipment	20
1620	1640	Completed survey	Data collection	20
1640	1705	Cleanup.	Daily close-up	25

Company: IT Jewell Date: 8 March 2006			**Personnel: Shawn Jewell, Peter Cooke**	
Start	**Stop**	**Remarks**	**Activity**	**Chargeable**
0730	0830	ATC delay - boat motor repaired	Downtime (ATC)	60
0900	1030	Mounted mini-sensor	Daily setup	90
1030	1100	Perimeter survey	Data collection	30
1100	1115	Dressed warmer for main survey	Downtime	15
1115	1215	Surveyed using a criss-cross pattern, minor axis of pond.	Data collection	60
1215	1300	Lunch	Nonchargeable downtime	45
1300	1510	Completed survey	Data collection	130
1510	1600	Packed up	Daily close-up	50

Company: IT Jewell Date: 9 March 2006			**Personnel: Shawn Jewell, Peter Cooke**	
Start	**Stop**	**Remarks**	**Activity**	**Chargeable**
0730	1205	Loaded sonar equipment and transported to ATC	Initial setup	275
1205	1220	Lunch	Nonchargeable downtime	15
1220	1355	Attached Odum sonar unit to sensor sled. Sonar head mounted at the same depth as magnetometers	Daily setup	95
1355	1515	Sonar mapped the pond - completed.	Data collection	80
1515	1540	Mounted G-882 on sled and instrumented boat. (A calibration run using the G-882 sensor was needed to support the data analysis process.)	Calibration	25
1540	1650	Resurveyed turn locations to reduce noise produced when boat turned.	Data collection	70
1650	1725	Cleanup	Daily close-up	35

APPENDIX C. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced ordnance item.

Detection: An anomaly location that is within R_{halo} of an emplaced ordnance item.

Munitions and Explosives Of Concern (MEC): Specific categories of military munitions that may pose unique explosive safety risks, including UXO as defined in 10 USC 101(e)(5), DMM as defined in 10 USC 2710(e)(2) and/or munitions constituents (e.g. TNT, RDX) as defined in 10 USC 2710(e)(3) that are present in high enough concentrations to pose an explosive hazard.

Emplaced Ordnance: An ordnance item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., non-ordnance item) buried by the government at a specified location in the test site.

R_{halo}: A pre-determined radius about the periphery of an emplaced item (clutter or ordnance) within which a location identified by the demonstrator as being of interest is considered to be a response from that item. For the purpose of this program, a circular halo 0.5 meters in radius will be placed around the center of the object for all clutter and ordnance items less than 0.6 meters in length. When ordnance items are longer than 0.6 meters, the halo becomes an ellipse where the minor axis remains 1 meter and the major axis is equal to the projected length of the ordnance onto the ground plane plus 1 meter.

Response Stage Noise Level: The level that represents the point below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the Blind Grid test area.

Discrimination Stage Threshold: The demonstrator selects the threshold level that they believe provides optimum performance of the system by retaining all detectable ordnance and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for n independent trials with the probability p of success and the probability 1-p of failure being the same for each trial. The number of successes x observed in the n trials is an estimate of p and is considered to be a binomially distributed random variable.

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/(No. of emplaced ordnance in the test site).

Response Stage False Positive (fp^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of False Positive (P_{fp}^{res}): P_{fp}^{res} = (No. of response-stage false positives)/(No. of emplaced clutter items).

Response Stage Background Alarm: An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open water or littoral scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind Grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open water only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{fp}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can, therefore, be written as $P_d^{res}(t^{res})$, $P_{fp}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to response-stage data that discriminates ordnance from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to ordnance, as well as those that the demonstrator has high confidence correspond to non-ordnance or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced ordnance in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm: An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open water or littoral scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can, therefore, be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] Figure A-1 shows how P_d versus P_{fp} and P_d versus BAR are combined into ROC curves. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

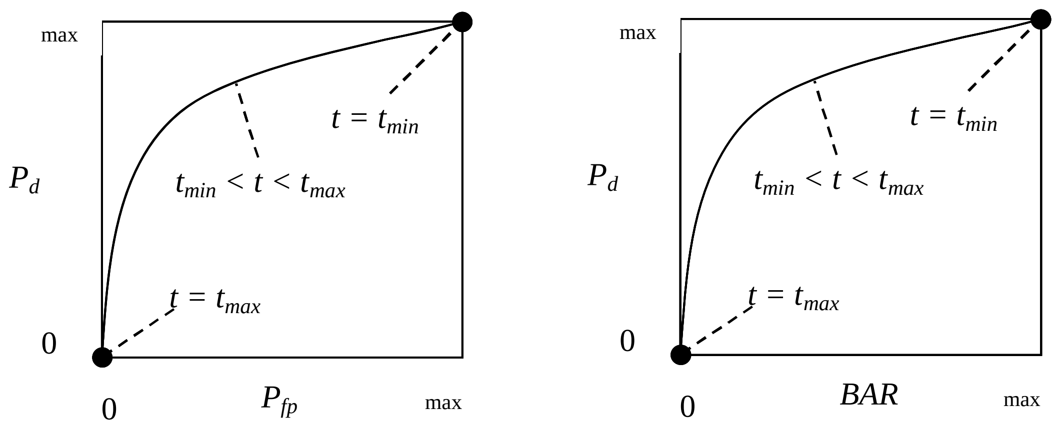

Figure A-1. ROC curves for open-site testing. Each curve applies to both the response and discrimination stages.

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a predetermined and fixed number of detection opportunities (some of the opportunities are located over ordnance and others are located over clutter or blank spots). In an Open Water scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the Open Water ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the Blind Grid test sites are true ROC curves.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. The efficiency measures the amount of detected ordnance retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$: measures (at a threshold of interest), the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage t_{min}) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the ordnance initially detected in the response stage was retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{fp}^{res}(t_{min}^{res})]$: measures (at a threshold of interest), the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage t_{min}). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind Grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$
Open water: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})])$

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON EXPLANATION:

The Chi-square test for differences in probabilities (or 2 x 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations (ref 4, pages 144 through 151).

A one-sided 2 x 2 contingency table is used in the Shallow Water Site Program to compare each area (Open Water, Littoral, Deep Water) to the Blind Grid since each area introduces a water feature that makes it potentially more difficult to survey than the Blind Grid. The one-sided 2 x 2 contingency table is used to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X's system is significantly degraded by the more challenging feature introduced. A two-sided 2 x 2 contingency table is used to compare performance between any two of the test sites other than the Blind Grid, to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X's system is significantly different between those two test sites.

The test statistic of the 2 x 2 contingency table is the Chi-square distribution with one degree of freedom. For the one-sided test, a significance level of 0.05 is chosen which sets a critical decision limit of 3.84 from the Chi-square distribution with one degree of freedom. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's Exact Test is used and the critical decision limit is the chosen significance level, which is 0.05 for one-sided tests and 0.10 for two-sided tests. With Fischer's test, if the test statistic (p-value) is less than the critical value, then the null hypothesis of similar performance is rejected in favor of the alternative hypothesis: significantly greater than for the one-sided case or significantly different for the two-sided case.

Shallow-water UXO Detection Test Site examples, where blind grid results are compared to those from the open water and littoral sites and the non-grid sites (open water and littoral) are compared to each other as follows. It should be noted that a significant result does not prove a cause and effect relationship exists between the change in survey area and sensor performance; however, it does serve as a tool to indicate that one data set reflects relatively degraded system performance of a large enough scale than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying each of the three areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind Grid	Open water	Littoral
P_d^{res}	100/100 = 1.0	8/10 = .80	20/33 = .61
P_d^{disc}	80/100 = 0.80	6/10 = .60	8/33 = .24

P_d^{res}: BLIND GRID versus OPEN WATER. Using the example data above to compare probabilities of detection in the response stage, all 100 ordnance out of 100 emplaced ordnance items were detected in the blind grid while 8 ordnance out of 10 emplaced were detected in the open water. Fischer's test must be used since a 100 percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic (p-value) of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open water relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN WATER. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 emplaced ordnance items were correctly discriminated as ordnance in blind grid testing while 6 out of 10 emplaced ordnance items were correctly discriminated as such in open water testing. Those four values are used in the Chi-square Contingency Test to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 3.84, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: BLIND GRID versus LITTORAL. Using the example data above to compare probabilities of detection in the response stage, 100 out of 100 and 20 out of 33 are used to calculate a test statistic (< 0.000) that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.61) is considered to be significantly less at the 0.05 level of significance.

P_d^{disc}: BLIND GRID versus LITTORAL. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 and 8 out of 33 emplaced ordnance items were correctly discriminated as such in open water testing. Those four values are used to calculate a test statistic of 32.01. Since the test statistic is greater than the critical value of 3.84, the smaller discrimination stage detection rate (0.24) is considered to be significantly less at the 0.05 level of significance.

P_d^{res}: OPEN WATER versus LITTORAL. Using the example data above to compare probabilities of detection in the response stage, 8 out of 10 and 20 out of 33 are used to calculate a test statistic of 0.56. Since the test statistic is less than the critical value of 2.71, the two response stage detection rates are considered to be not significantly different at the 0.10 level of significance.

P_d^{disc}: OPEN WATER versus LITTORAL. Using the example data above to compare probabilities of detection in the discrimination stage, 6 out of 10 and 8 out of 33 are used to calculate a test statistic of 2.98. Since the test statistic is greater than the critical value of 2.71, the two discrimination stage detection rates are considered to be significantly different at the 0.10 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and change in performance, it does indicate that the ability of Demonstrator X to correctly discriminate seems to have been degraded by features of the littoral area relative to results from the open water using the same system.

APPENDIX D. REFERENCES

1. Environmental Quality Technology - Operational Requirements Document (EQT-ORD) for: A(1.6.a): UXO Screening, Detection and Discrimination.

2. Proposal for Shallow Water Unexploded Ordnance (UXO) Detection & Discrimination Technology Demonstration Volume 1 - Technical/Management (Plan). Submitted in response to BAA W91ZLK-04-R-0001, by IT Jewell, 30 August 2005.

3. IT Jewell Final Report, Geometrics G-882 and MM Explorer Mini Magnetometers Shallow Water Unexploded Ordnance (UXO) Detection & Discrimination Technology Demonstration, 10 April 2006.

4. Practical Nonparametric Statistics, W.J. Conover, John Wiley & Sons, 1980, pages 144 through 151.

APPENDIX E. ABBREVIATIONS

ADST = Aberdeen Data Services Team
APG = Aberdeen Proving Ground
ATC = U.S. Army Aberdeen Test Center
BAA = Broad Agency Announcement
BAR = background alarm rate
DMM = discarded military munitions
EQT = Army Environmental Quality Technology Program
EQT-ORD = Environmental Quality Technology - Operational Requirements Document
ERDC = U.S. Army Corps of Engineers Engineering, Research and Development Center
ESTCP = Environmental Security Technology Certification Program
GPS = Global Positioning System
LED = light-emitting diode
MEC = munitions and explosives of concern
MEDTC = Military Environmental Technology Demonstration Center
P_{ba} = probability of background alarm rate
P_d = probability of detection
P_d^{disc} = probability of detection, discrimination stage
P_d^{res} = probability of detection, response stage
P_{fp} = probability of false positive
P_{fp}^{disc} = probability of false positive, discrimination stage
P_{fp}^{res} = probability of false positive, response stage
POC = point of contact
QA = quality assurance
QC = quality control
ROC = receiver operating characteristics
SERDP = Strategic Environmental Research and Development Program
TMI = Total Magnetic Intensity
USAEC = U.S. Army Environmental Command
UXO = unexploded ordnance

APPENDIX F. DISTRIBUTION LIST

DTC Project No. 8-CO-160-UXO-016

Addressee	No. of Copies
Commander U.S. Army Aberdeen Test Center ATTN: CSTE-DTC-AT-SL-E (Mr. Gary Rowe) (Library) CSTE-DTC-AT-CS-R Aberdeen Proving Ground, MD 21005-5055	 1 1 1
Booz Allen Hamilton ATTN: (Mr. Patrick McDonnell) 4692 Millennium Drive, Suite 200 Belcamp, MD 21017-1535	 1
IT Jewell, Inc. ATTN: (Mr. Shawn P. Jewell) 11654 Plaza America Drive No. 631 Reston, VA 20190-4700	 1
Defense Technical Information Center 8725 John J. Kingman Road, Suite 0944 Fort Belvoir, VA 22060-6218	 PDF

Secondary, “hard-copy” distribution is controlled by Commander, U.S. Army Environmental Command, ATTN: SFIM-AEC-ATT.